

A feature-tracking model simulates the motion direction bias induced by phase congruency

M. Michela Del Viva

Dipartimento di Psicologia, Università di Firenze,
Firenze, Italy, & Istituto di Neuroscienze CNR, Pisa, Italy



M. Concetta Morrone

Università Vita-Salute S Raffaele, Milano, Italy, &
Istituto di Neuroscience CNR, Pisa, Italy



Here we report a new motion illusion where the prevailing motion direction is strongly influenced by the relative phase of the harmonic components of the stimulus. The basic stimulus is the sum of three sinusoidal contrast-reversing gratings: the first, the third, and the fifth harmonic of two square wave gratings that drift in opposite direction. The phase of one of the fifth components was kept constant at 180 deg, whereas the phase of the other fifth harmonic was varied over the range 0–150 deg. For each phase value of the fifth harmonic, the motion was strongly biased toward its direction, corresponding to the direction with stronger phase congruency between the three harmonics. The strength of the prevailing motion was assessed by measuring motion direction discrimination thresholds, by varying the contrast of the third and the fifth harmonics plaid pattern. Results show that the contrast of high harmonics had to be increased by more than a factor of 10, to achieve a balance of motion for phase differences greater than 60 deg between the 2 fifth harmonics. We also measured the dependence on the absolute phase of harmonic components and found that it is not an important parameter, excluding the possibility that local luminance cues could be mediating the effect.

A feature-tracking model based on previous work is proposed to simulate the data. The model computes local energy function from a pair of space-time separable front stage filters and applies a battery of directional second stage mechanisms. It is able to simulate quantitatively the phase congruency dependence illusion and the insensitivity to overall phase. Other energy models based on directional filters fail to simulate the phase congruency dependency effect.

Keywords: motion, transparency, phase congruency, local energy model, feature tracking, second order motion, energy models

Introduction

There are several examples of moving visual stimuli comprising a small number of harmonic components that produce interesting phenomena: the components can be perceived as to move independently or to form a coherent moving pattern depending on the particular parameters of each component.

A nice example, easy to replicate in any laboratory, is provided by two sinusoidal gratings of equal contrast, orientation and spatial frequency moving with the same speed in opposite directions. The stimulus appears as a single grating sinusoidally modulated in contrast (counter-phase grating) over a wide range of spatiotemporal frequencies and contrasts (see [Figure 1A](#) and [Movie 1](#)). Conversely, two gratings with square wave luminance profiles drifting in opposite directions, which mathematically correspond to a sum of counterphase sinusoidal gratings of different frequencies, are perceived as two distinct patterns drifting in transparency one over the other (see [Figure 1B](#) and [Movie 2](#)), especially if presented gradually over time. Intermediate situations, such as two pairs of gratings contrast reversing sinusoidally at different frequencies, may lead to an ambiguous and alternate per-

ceptions of flicker and transparent motion (see [Figure 1C](#) and [Movie 3](#)), as observed for example for the first and the third harmonic of a square wave. Visual perception changes gradually from flicker to transparent motion, by adding flickering sinusoidal gratings of increasing frequencies ([Figures 1C](#) and [B](#); [Movies 3](#) and [2](#)).

These observations on the appearance of simple moving stimuli form part of a more general issue of image segmentation:

- How are different moving components of objects grouped together in the visual scene?
- Are they grouped according to their spatial frequency content or is image segmentation based on grouping elements with similar velocity?
- Is segmentation achieved at an early stage of visual analysis or does it require a priori knowledge about the physical world and a top-down modulation from higher cognitive processes?

It seems that in the first example mentioned above ([Movie 1](#)), the two gratings are “grouped” together according to their spatial frequency content because no net motion in any direction is perceived (as shown by the blue ellipse of [Figure 1A](#), right), whereas in the second example ([Movie 2](#)) components with the same direction are

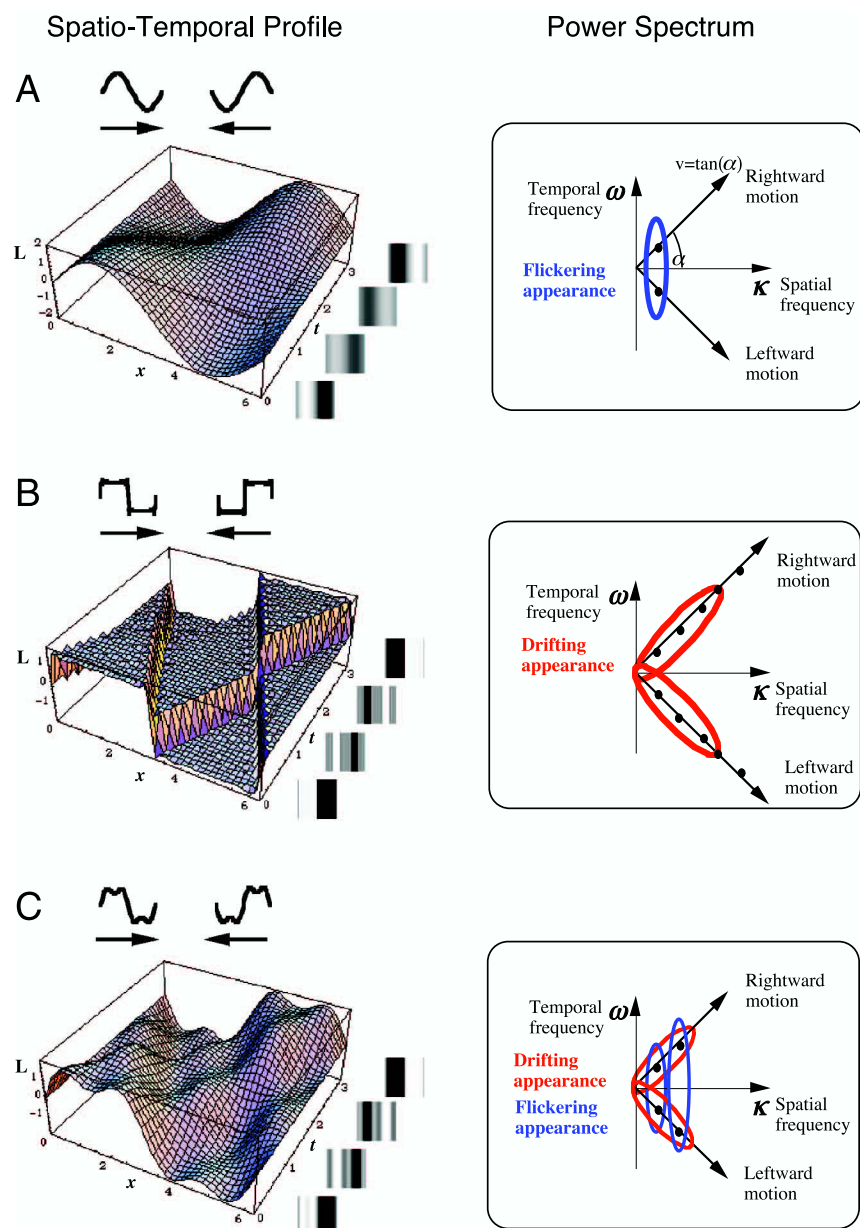


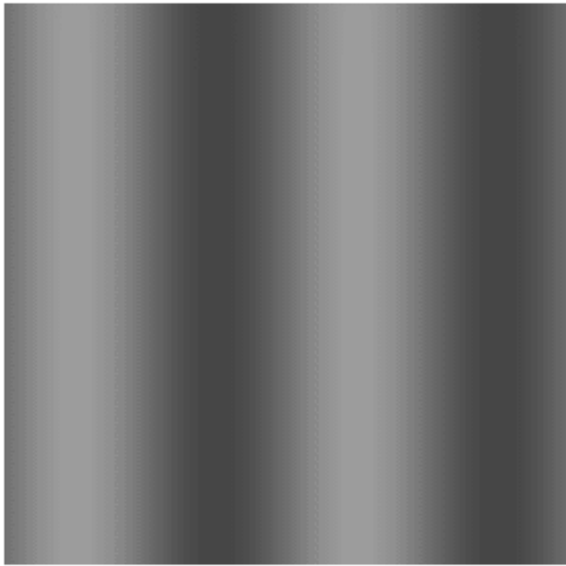
Figure 1. Perceptually different visual motion of stimuli with similar power spectra: (A) Two identical sinusoidal gratings moving in opposite directions. (B) Two identical square wave gratings moving in opposite directions. (C) Two pairs of sinusoidal gratings (first and third harmonics of the square wave grating) moving in opposite directions. Left: spatiotemporal luminance profiles for the x - t plane [$L(y)$ is constant]. Right: Fourier power spectra of the stimuli. Red ellipses collect components according to the velocity of the pattern; blue ellipses collect components according to their spatial frequency (sinusoidally contrast modulated gratings).

grouped together (as shown by the two red ellipses of Figure 1B, right).

The goal of this paper is to study quantitatively the transition between perception of flicker and of transparent motion and to simulate both perceptual effects using a single motion sensitive mechanism. In particular, we will show that the critical parameter for the transition from flicker to transparent motion is the relative phase (phase congruency) between the spatiotemporal Fourier components. We will also show that the data can be well simulated by a local energy model extended to the temporal domain.

For static images, it is known that recognition of the image depends on relative phase of Fourier components (Openheim & Lim, 1981; Piotrowski & Campbell, 1982) and that the perceptual structure of the images in salient features is dictated by the organization of the peak of local energy function (Morrone & Burr, 1988). By definition, local energy function is highly sensitive to the local phase congruency of various harmonic components and can predict quantitatively transparency effects of still images.

For motion, we have proposed an algorithm that, similarly to earlier edge-detection-tracking models (Hildreth,

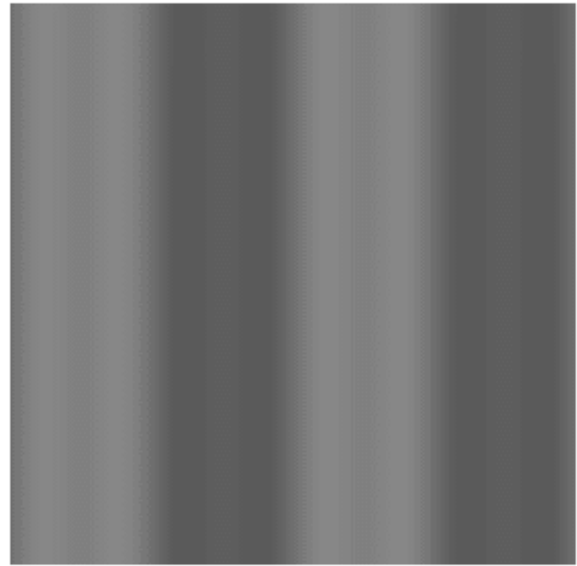


Movie 1. Sinusoidal grating whose contrast is sinusoidally modulated over time. The luminance profile and the amplitude spectra are shown in [Figure 1A](#).

1984; Marr & Ullman, 1981), first extracts visually salient features and then computes their velocity by tracking their energy over time. The feature-tracking approach to the analysis of motion has been somewhat neglected over the last few decades, being considered biological implausible. However, some experimental evidence suggests that feature-tracking may be used by the human visual system (Cavanagh & Mather, 1989; Derrington & Ukkonen, 1999; Georgeson & Shackleton, 1989; Morgan, 1992;



Movie 2. Two identical square wave gratings drifting in opposite direction. The luminance profile and the amplitude spectra are shown in [Figure 1B](#).



Movie 3. Two pairs of sinusoidal gratings whose contrast is sinusoidally modulated over time, corresponding to the first and the third harmonics of the drifting square wave stimulus of [Movie 2](#). The luminance profile and the amplitude spectra are shown in [Figure 1C](#).

Morgan & Mather, 1994; Pantle & Turano, 1992; Seiffert & Cavanagh, 1998). In particular, using the pedestal paradigm developed by Lu and Sperling, it is possible to bias the direction of a compound grating towards the trajectory of features (Lu & Sperling, 1995).

However, the prevailing direction is more often biased towards the direction of the prevailing energy of the stimuli (Georgeson & Scott-Samuel, 1999; Zaidi & DeBonet, 2000).

The perception of motion transparency elicited by superimposed random-dot fields (Qian, Andersen, & Adelson, 1994a) or by superposed sinusoidal components of different orientations (plaids; Wilson, Ferrera, & Yo, 1992) can be simulated by motion energy models. These models compute motion energy independently for a battery of directional filters, tuned to various velocities and directions, and apply an opponent stage between opposite directions to compute locally or globally the net motion. Here we use three counterphase gratings close in the spatiotemporal domain so as to excite preferentially a stationary nondirectional filter, with balanced energy in the two directions of motion. We will show that energy models based on directional filters cannot mediate the capture effect of phase congruency. On the other hand, a model that shares many similar characteristics to the original model proposed by Chubb and Sperling (1988) for the detection of second order motion and to a subsequent extension for the detection of third order motion (Lu & Sperling, 2001, 2002) can simulate quantitatively the data. To simulate the capture on motion direction by phase congruency, it is necessary to use a front stage mechanism that computes local energy function followed by a directional second stage mechanism (Chubb & Sperling, 1988).

Experimental methods

Stimuli

Stimuli were generated frame by frame on a Silicon Graphics Iris-35 workstation using dedicated software and the HIPS image processing package and displayed on a Barco Calibrator monitor (CDCT 6551) at 120 Hz temporal resolution. The monitor was driven by a graphic board (Cambridge Res. System 15 bits resolution) under the control of a PC computer. The whole stimulus subtended 20 deg of visual angle at a viewing distance of 57 cm.

Mean luminance, measured with a digital photometer Milta CS100, was 30 cd/m² and luminance nonlinearities were corrected separately for each of the three guns of the monitor to balance the chromaticity on the whole range of luminance.

The luminance profile of the stimulus used in first experiment, expressed as a Fourier series, is described by Equation 1,

$$\begin{aligned}
 I(x, t) = & L_0(x, t) \left(1 + C [\sin(\kappa_0 x + \omega_0 t) + \sin(\kappa_0 x - \omega_0 t)] \right) \\
 & + CA \frac{1}{3} [\sin(3\kappa_0 x + 3\omega_0 t) + \sin(3\kappa_0 x - 3\omega_0 t)] \\
 & + CA \frac{1}{5} [\sin(5\kappa_0 x + 5\omega_0 t + \phi) + \sin(5\kappa_0 x - 5\omega_0 t + \pi)].
 \end{aligned}
 \tag{1}$$

The power spectrum is schematically shown in Figure 2A. The luminance profile in the vertical spatial dimension is constant. The w_0 and k_0 are the temporal and spatial frequencies of the fundamental equal to 2 Hz and 0.5 c/deg. C is the contrast of the fundamental and A the ratio of the contrast of the high frequency plaid pattern to the contrast of the fundamentals.

Note that the 2 fifth harmonics differ in phase by $\phi - \pi$. The shift by π of one of the fifth harmonics was necessary to decrease the saliency of motion in that direction so it could be balanced by increasing the contrast of the high frequency plaid pattern in the motion discrimination task (see detailed explanation in the Experimental results section).

The luminance profile of the stimulus used in the second experiment, expressed as a Fourier series, is described by Equation 2,

$$\begin{aligned}
 I(x, t) = & L_0(x, t) \left(1 + C [\sin(\kappa_0 x + \omega_0 t + \psi) + \sin(\kappa_0 x - \omega_0 t + \psi)] \right) \\
 & + CA \frac{1}{3} [\sin(3\kappa_0 x + 3\omega_0 t + \psi) + \sin(3\kappa_0 x - 3\omega_0 t + \psi)] \\
 & + CA \frac{1}{5} [\sin(5\kappa_0 x + 5\omega_0 t + \phi) + \sin(5\kappa_0 x - 5\omega_0 t + \psi + \pi)].
 \end{aligned}
 \tag{2}$$

The power spectrum is schematically shown in Figure 2B. This stimulus also has a constant luminance profile in the vertical spatial dimension.

Note that here ψ represents the absolute phase of the stimulus. Again 1 fifth harmonic is phase shifted of π (see detailed explanation in the Experimental results section).

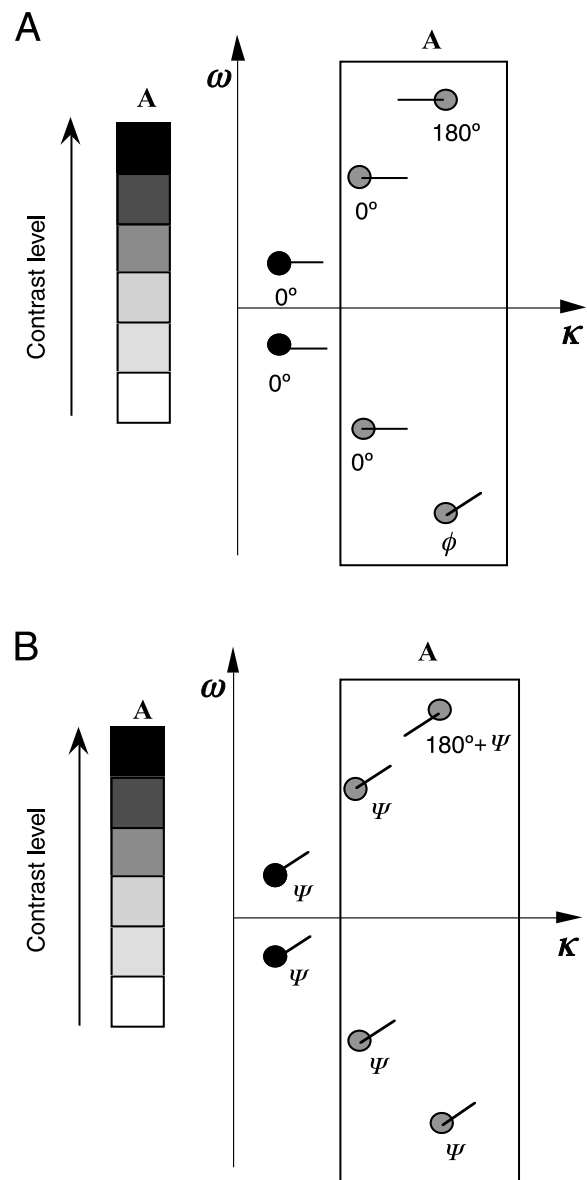


Figure 2. Schematic representation in the frequency domain of the stimuli and procedure. (A) Experiment 1: direction discrimination of prevailing motion as a function of contrast (parameter A) of high frequency components for different values of phase ϕ of 1 fifth harmonic. The phase of the other fifth harmonic has a constant phase shift of π . The contrast of fundamentals is kept constant. (B) Experiment 2: direction discrimination of motion of features as a function of contrast (parameter A) of high frequency components for different values of absolute phase ψ . The phase of the other fifth harmonic has an additional phase shift of π . The contrast of fundamentals is kept constant.

Procedure

In both experiments, we measured the threshold for motion direction discrimination as a function of the contrast of higher frequency plaid pattern (parameter A in Equations 1 and 2, and boxes in Figure 2). This is an

evaluation of the transition point between perception of flicker and perception of directional motion.

We measured the influence on the perception of motion of relative (ϕ , first experiment) and of absolute (ψ , second experiment) phase. In particular, for any fixed value of phase, we measured the contrast of the higher frequency components (parameter A in Equations 1 and 2; see box in Figure 2) to perceive with equal probability leftward and rightward motion. Contrast of each fundamental frequency (parameter C in Equations 1 and 2) remained constant and set to 0.1 in the first experiment and to 0.05, 0.1, and 0.3 in the second experiment (see Figure 2). The contrast A was varied adaptively with the QUEST staircase algorithm (Watson & Pelli, 1983). Subjects reported, in a single-interval 2AFC procedure with feedback, the perceived direction of motion by pressing a CB1 button (Cambridge Research Systems). Correct motion direction was assigned arbitrarily to the direction of the fifth harmonic with variable phase. For any condition tested, data were obtained with more than five QUEST staircases, each comprising more than 40 trials.

The stimulus presentation duration was 1 sec, vignetted with a Gaussian temporal envelope of time constant 250 ms to avoid transients at stimulus onset and offset. Values of temporal frequency and exposure were such that one entire period of fundamental was displayed to avoid spectral distortions of the original stimulus.

Two subjects participated to the experiment, one of the authors and the other naive to the goals of the experiment.

For each condition and for each subject, a cumulative maximum likelihood fit was performed off-line with all data, obtained in all sessions, using a Weibull psychometric function, described by Equation 3,

$$P(i) = 1 - 0.5e^{-\beta(x_i - T)}. \quad (3)$$

This curve expresses, on a logarithmic scale, the probability of motion direction discrimination in a 2AFC task as a function of contrast x_i .

The fitting procedure had one free parameter T ; β —representing the slope of the psychometric curves—was set equal to 1.75. This value was assessed by using a two-parameter fit of the psychometric curves and by taking the average value across conditions and subjects. The variation of β across conditions was not statistically significant.

Contrast threshold T was defined as the contrast corresponding to 0.75 of the fitted curve. Threshold measurements were repeated for different values of phase shift ranging from 0 to $5\pi/6$. A two-interval 2AFC was used to measure contrast detection threshold of the higher frequency components, using the same analysis and fitting procedure as for the motion discrimination threshold, with β set to 1.75.

Experimental results

Experiment 1: Dependency on relative phase of components

The basic stimulus was the sum of three sinusoidal contrast-reversing gratings, the first, the third, and the fifth harmonics of two drifting square wave gratings (see Figure 2A). To study the effect of relative phase, the phase of one of the fifth components was altered and was kept constant at 180 deg. The motivation for this phase manipulation becomes clear by observing Figure 3 that shows the spatiotemporal profiles of two example stimuli. The stimulus in Figure 3A has the phase of both fifth components equal to zero (Figure 3A), the stimulus in Figure 3B has 1 fifth component with phase equal to zero, the other with phase equal to 180 deg (corresponding to Equation 1 with $\phi = 0$). When there is no phase shift (Figure 3A), two clear pairs of edges oriented at ± 45 deg are visible and have similar contrast: In this case, perception of direction of motion is ambiguous (see Movie 4). When the shift is 180 deg (Figure 3B), the pair of edges along 45 deg are weaker and less defined, given that the luminance profile along this velocity approximate a more triangular waveform as a consequence of the phase shift. The stimulus in Figure 3B elicits a clear motion perception bias in the direction opposite to the phase manipulation (see Movie 5).

The first experiment studies the effect of the relative phase between the fifth harmonics on motion direction bias. For a given value of ϕ , we vary the contrast of the third and the fifth harmonics to discriminate a global direction of motion. An example psychometric function is shown in Figure 4 for $\phi = 0$ deg (black curve). By increasing the contrast of higher harmonics (parameter A in Equation 1; see box in Figure 2), perception changes smoothly from flicker to transparent motion. When the contrast of higher harmonics is close to 0, the stimuli comprises only two sinusoidal gratings of the same amplitude, spatial frequency and speed, but opposite direction (components in the rectangle in Figure 2 are set to zero) leading to perception of flicker and to a chance performance. Conversely, when contrast of higher harmonics is set to maximum (0.03 for each third harmonic and 0.02 for each fifth harmonic), both directions of motion are perceived. This condition would still produce chance performance in a motion discrimination task. However, the phase shift of π of one of the fifth components dampens the saliency of motion in the direction of this harmonic component (see Movie 5) and the transition between transparency and flicker can be assessed by measuring direction discrimination of motion. The balance for the example in Figure 4 is obtained for $CA \approx 0.018$ for $\phi = 0$ deg (black psychometric curve) (Equation 1). However, the balance is

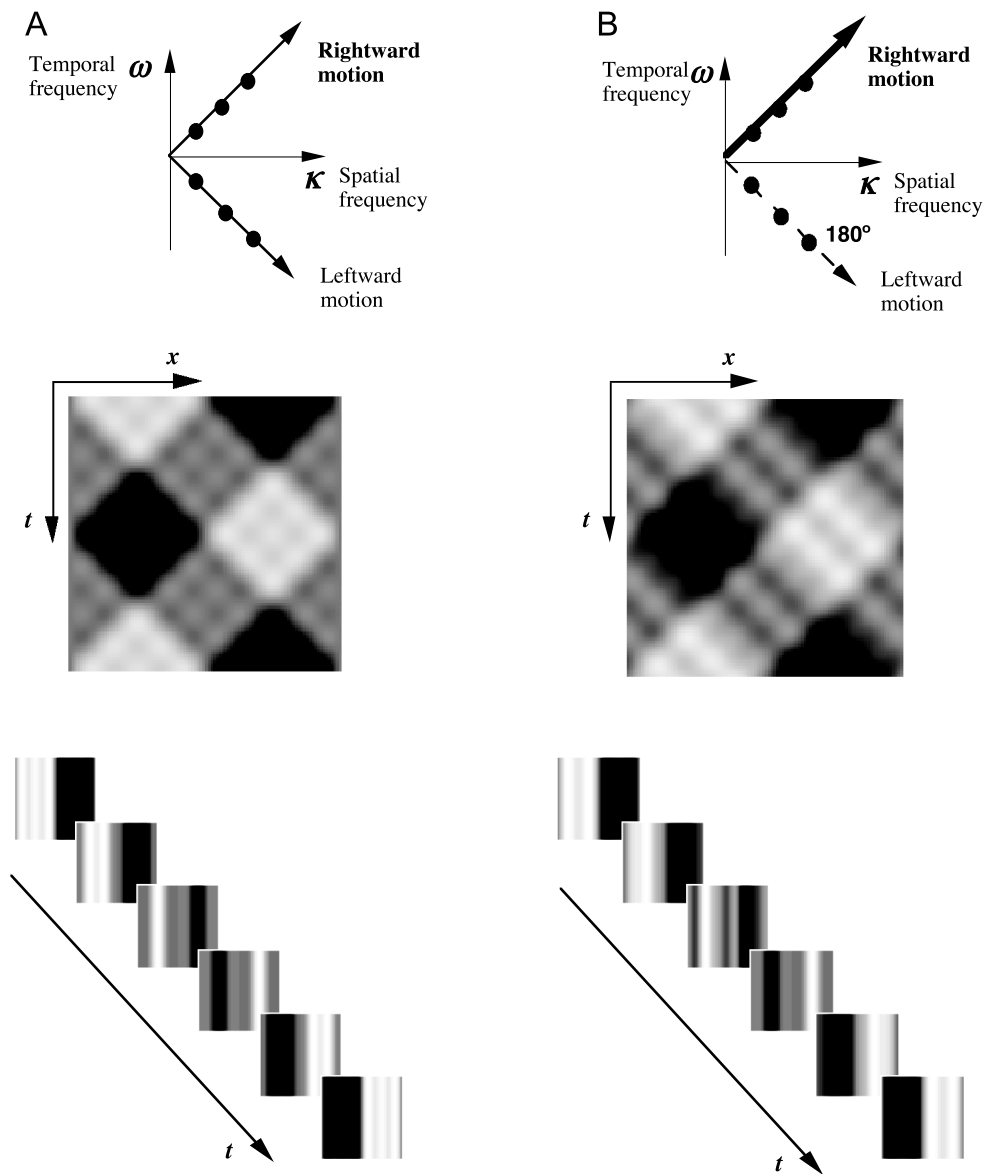


Figure 3. Perceptual effects of relative phase on image segmentation: (A) Three pairs of sinusoidal gratings moving in opposite directions, perceived as two gratings moving transparently. (B) After shifting by π , the phase of one component, the same stimulus is perceived as a single grating moving over a flickering pattern. Upper panels: Fourier power spectra. Middle panels: spatiotemporal representations of one period of the stimuli. Lower panels: temporal sequence of the stimuli.

achieved for $CA \approx 0.04$ when the phase $\phi = 120$ deg (red psychometric curve; [Movie 6](#)). To perceive a prevailing direction of motion, the subjects have to increase the contrast of the higher harmonic components by more than a factor of two when the phases of the 2 fifth components are more similar. We will refer in the following to the salient direction of motion as to “feature motion.” The justification of the use of this term will be given in the model section.

[Figure 5](#) shows sensitivity to feature motion as a function of phase shift (ϕ in [Equation 1](#)). The ordinate plots the inverse of the multiplication of the parameters C and A in [Equation 1](#). Quantitative results confirm previous qualitative observations. An increase in phase shift of the fifth harmonic (ϕ) decreases sensitivity to feature motion

in the direction of the component with phase ϕ . When ϕ is zero, a bias in motion direction is perceived at twice the detection threshold of the higher component (dashed curves of [Figure 5](#)). When $\phi = 130$ deg, more than a log unit of super-threshold contrast is needed to achieve a preference in motion direction, showing a strong dependence of motion upon the phase congruency between the various components. These results indicate that grouping across the same velocity components is enhanced when phase between harmonic is similar.

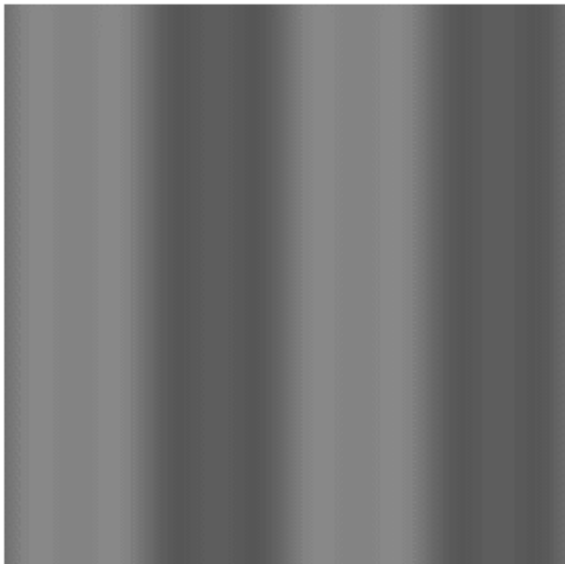
The results are highly consistent between the two subjects. Note also that phase dependency is very sharp especially for phase shift over $\pi/2$. This means that sensitivity to feature motion is particularly low for phase shifts over $\pi/2$.



Movie 4. Three pairs of sinusoidal gratings, whose contrast is sinusoidally modulated over time, corresponding to the first, the third, and the fifth harmonics of the drifting square wave stimulus of [Movie 2](#). The luminance profile and the amplitude spectra are shown in [Figure 3A](#).

Experiment 2: Dependency on absolute phase of components

[Figure 6](#) shows sensitivity as a function of absolute phase shift (ψ in [Equation 2](#)). Changing absolute phase changes dramatically the spatiotemporal luminance profile of the stimuli. For example, the prevailing edges of



Movie 5. Three pairs of sinusoidal gratings, whose contrast is sinusoidally modulated over time as for the stimulus of [Movie 4](#), except than the leftward fifth harmonic is phase shifted by π . The luminance profile and the amplitude spectra are shown in [Figure 3B](#).

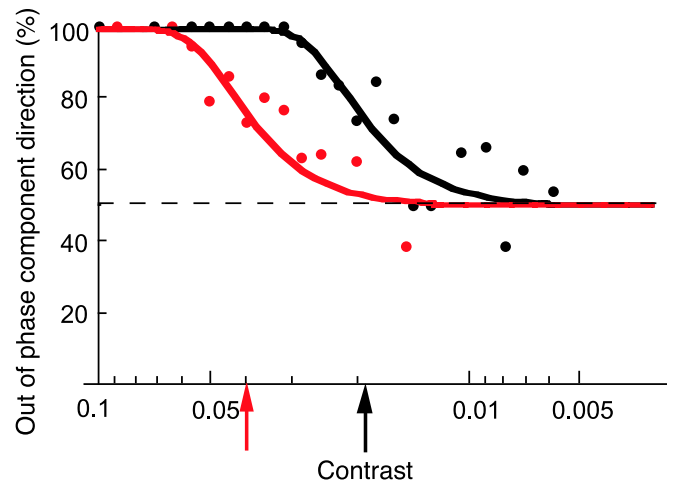


Figure 4. Psychometric functions of one subject—to perceive the direction of motion of the variable phase fifth component, whose phase was set to 0 deg for the black curve and to 120 deg for the red curve. The abscissa plots the ratio of the contrasts of the plaid stimuli comprising all higher harmonic to the contrast of the fundamentals, given by the parameter A of [Equation 1](#). Each point is the average of at least 20 trials.

[Figure 3A](#) are transformed into lines when $\psi = 90$ deg. However, there is almost no dependency of motion perception on phase shift. All subjects reported that when the phase offset was varied, the prevailing motion was attributed to different kinds of features, such as moving square wave gratings or sawtooth gratings. Nevertheless, as soon as the higher harmonics were perceived, the perception of flicker broke down to transparent motion. We measured the effect for three base values of the contrast of the fundamental harmonic. At higher contrasts, the sensitivity seems lower for phase offsets around 90 deg, suggesting a preference for edges than lines. However, the effect is very small, about a factor of 1.4.

Modeling sensitivity to phase

The psychophysical results of this paper show a strong dependency of perception of motion transparency on relative phase of harmonic components of one-dimensional gratings. The overall power of the left and the rightward motion is always balanced, so a simple model based on an overall energy measure ([Adelson & Bergen, 1985](#); [Grzywacz & Yuille, 1990](#); [Heeger, 1987](#)) would fail to predict qualitatively the results. However, models that are highly sensitive to phase congruency, such as the local energy model for spatial vision, or a general feature-tracking motion models are likely candidates to simulate the effect. Here we used an extension in the spatiotemporal domain of the local energy model for feature detection ([Del Viva &](#)



Movie 6. Three pairs of sinusoidal gratings whose contrast is sinusoidally modulated over time as for the stimulus of [Movie 3](#), except than the leftward fifth harmonic is phase shifted by π and the right ward fifth harmonic is phase shifted by 120 deg. For this stimulus, the two directions of motion are nearly balanced: $CA \approx 0.04$.

Morrone, 1998). Similar to earlier edge-detection-tracking models (Hildreth, 1984; Marr & Ullman, 1981), the model first extracts visually salient features and then computes the prevailing motion applying a second stage of analysis with operators that are velocity tuned.

Algorithm description

As in the original implementation of the model, the local energy function $E(x,t)$ in any point of the image is computed by convoluting the image $I(x,t)$ with pairs of band-pass spatial linear filters in quadrature phase $F_e(x,t)$ and $F_o(x,t)$ (Figure 7A, Equations 5 and 6). Local energy is then computed by summing the square of the outputs of convolution with each filter (Figure 7B, Equation 7). The local energy function is particularly sensitive to phase congruency: When the phases between harmonics are most similar, all the energy becomes concentrated in high peaks; for low phase congruency, the peaks become smoother and less defined. The local maxima correspond to the location of salient features. In the present algorithm (Figure 7), we did not segment the image by marking the features, but we use the local energy stage to transform the input in a function whose intensity is proportionally related to the salience of the spatial structure.

The obtained spatial saliency map was used to derive the prevailing motion direction. If the same feature travels at constant speed, the saliency map would show a single ridge

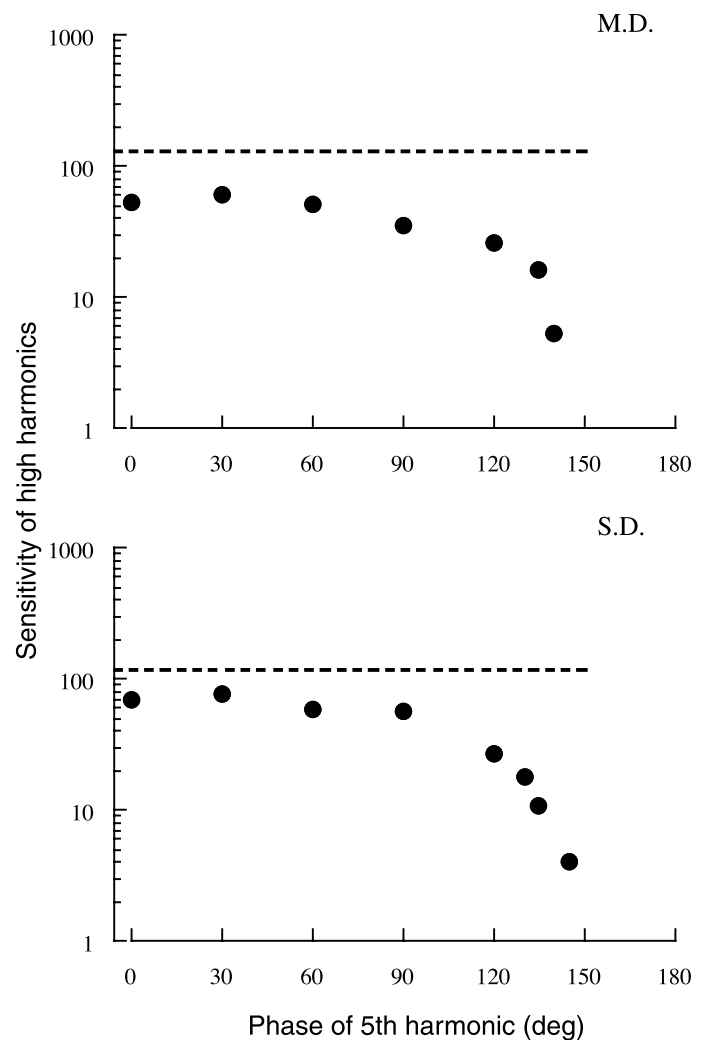


Figure 5. Contrast sensitivity of higher harmonics for motion direction discrimination, as a function of relative phase ϕ of the fifth harmonic, for two subjects. The sensitivity is given by $1/(AC)$ of Equation 1. Dotted lines represent contrast sensitivities for detection of the high harmonics compound.

in space-time and its orientation would give the velocity of the feature. An appropriate way to determine the orientation of the ridges is to convolve the saliency map again with detectors of various orientations and search for the detector that respond maximally.

For this second stage of analysis, we used oriented spatiotemporal filters tuned to different velocities. We evaluated several shapes of this battery of filters independently. Given that the stimulus comprises only two velocities, the local energy map obtained at first stage will excite maximally the second stage filter tuned at 45 and -45 deg (corresponding to velocity of ± 4 deg/s). This fact, which was also verified experimentally, allowed us to simplify further the model by measuring only the response from these two orientations (see Equations 8 and 9). To assess the prevailing direction of motion, we measured the average

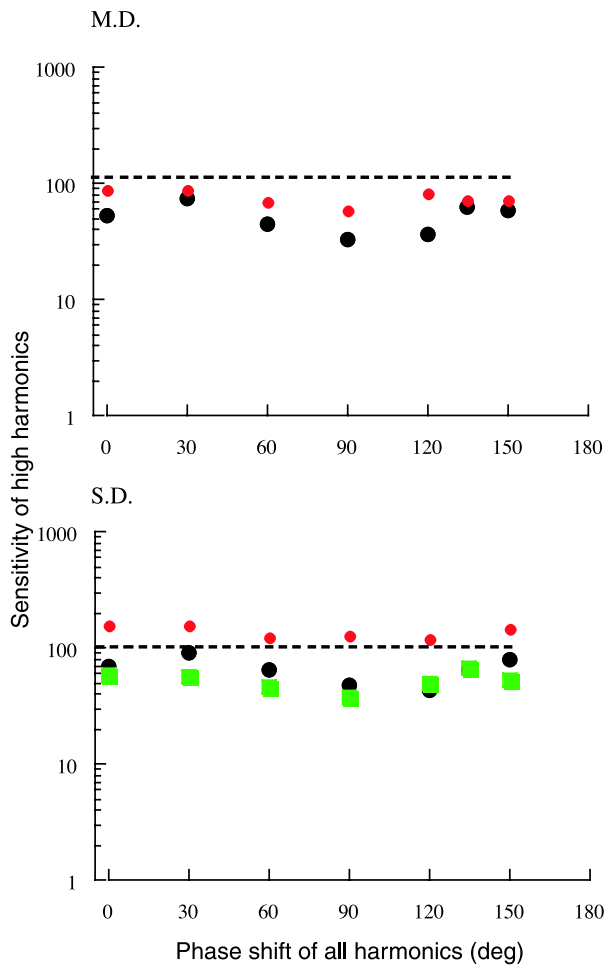


Figure 6. Contrast sensitivity of higher harmonics for motion direction discrimination as a function of relative phase ψ , with phase difference between the fifth components always equal to 180 deg. Different symbols represent different contrast of fundamentals (red diamonds: contrast = 0.05; black circles: contrast = 0.1; green squares: contrast = 0.3). Dotted line is the contrast sensitivity for detection of high harmonics when contrast of fundamentals is set to 0.1.

motion contrast from the outputs of the two velocity-tuned second/stage filters using the quantity:

$$O(\phi, K) = \frac{\int M_{\text{right}} dsdt - \int M_{\text{left}} dsdt}{\int M_{\text{right}} dsdt + \int M_{\text{left}} dsdt}, \quad (4)$$

where M_{left} and M_{right} are the outputs of convolution of the energy with filters (H_{left} and H_{right} in Equations 8 and 9) tuned to velocities at ± 4 deg/s (45 and -45 deg orientation in Figure 7C).

The quantity $O(\phi, K)$ can be thought as a measure of motion energy contrast and is obtained by integrating the output of the second stage filters over a full period

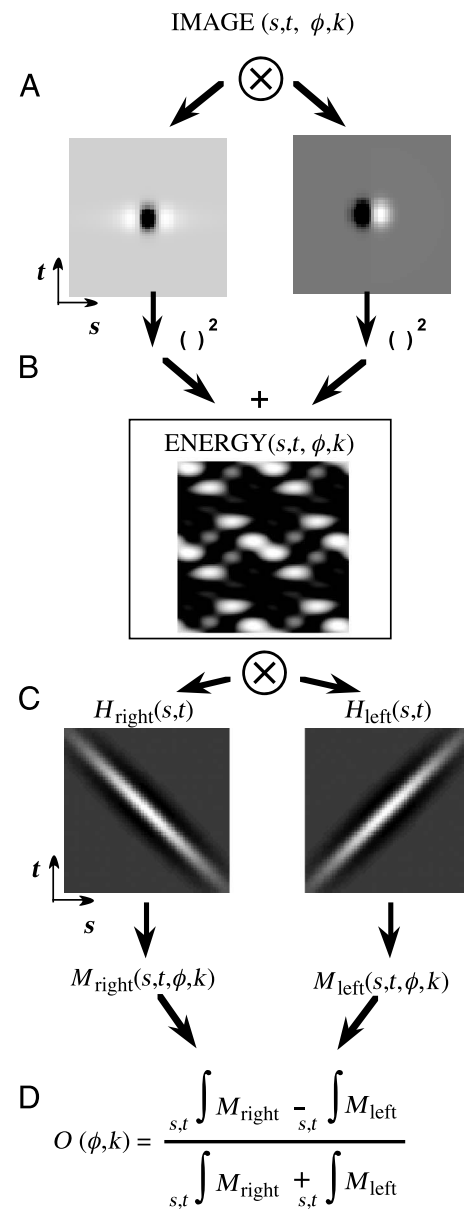


Figure 7. Different stages of the model. (A) Convolution with filters in quadrature phase oriented along the zero velocity; (B) computation of local energy by summing the square of the output of convolution with each filter; (C) convolution with second stage directional operator tuned at varies velocities; (D) equation used to evaluate the bias in motion direction from the overall response of the two most active velocity filters.

of the stimulus. It is important to note that the ratio $O(\phi, K)$ is not determined by the output of directional energy units, which is equivalent to applying a normalization to the output of motion opponent units, as previously used to simulate successfully the directional thresholds of two drifting grating (Georgeson & Scott-Samuel, 1999). We will show that the second stage filtering is essential to simulate the psychophysical results (see Figure 10).

The stimuli used in this experiment are very simple, close in spatiotemporal frequency domain and one dimensional, allowing a simplification at the implementation of the model. We used only one pair of filters and we left the actual shape (peak frequency and bandwidth) of the odd and even front-end filters as a free parameter. The orientation of the filters was always kept constant and tuned at zero velocity.

The output of the convolution of the input $I(x,t)$ with a generic even filter oriented along zero velocity, $F_e(x,t)$, is given by:

$$I(x,t) \otimes F_e(x,t) = C \left(\begin{aligned} &\cos(\omega_0 t + \kappa_0 x) - \cos(\omega_0 t - \kappa_0 x) \\ &+ \frac{-A \cos(5\omega_0 t - 5\kappa_0 x)}{5a_2} + \frac{A \cos(3\omega_0 t - 3\kappa_0 x)}{3a_1} \\ &+ \frac{A \cos(3\omega_0 t + 3\kappa_0 x)}{3a_1} + \frac{A \cos(5\omega_0 t + 5\kappa_0 x + \phi)}{5a_2} \end{aligned} \right). \quad (5)$$

The response of the odd filter is given by:

$$I(x,t) \otimes F_o(x,t) = C \left(\begin{aligned} &\sin(\omega_0 t - \kappa_0 x) + \sin(\omega_0 t + \kappa_0 x) \\ &+ \frac{A \sin(5\omega_0 t - 5\kappa_0 x)}{5a_2} - \frac{A \sin(3\omega_0 t - 3\kappa_0 x)}{3a_1} \\ &+ \frac{A \sin(3\omega_0 t + 3\kappa_0 x)}{3a_1} + \frac{A \sin(5\omega_0 t + 5\kappa_0 x + \phi)}{5a_2} \end{aligned} \right). \quad (6)$$

The shape of the front stage filters is given by the multiplicative constants a_1 and a_2 that represent the gain of the third and of the fifth harmonics respect to the fundamentals, respectively.

The energy is given by:

$$E(x,t) = (I(x,t) \otimes F_e(x,t))^2 + (I(x,t) \otimes F_o(x,t))^2 \\ = C^2 \left(\begin{aligned} &\left(\frac{-A \cos(5t-5x)}{5a_2} + \frac{A \cos(3t-3x)}{3a_1} - \cos(t-x) \right)^2 \\ &+ \cos(t+x) + \frac{A \cos(3t+3x)}{3a_1} + \frac{A \cos(5t+5x+\phi)}{5a_2} \end{aligned} \right)^2 \\ + C^2 \left(\begin{aligned} &\left(\frac{A \sin(5t-5x)}{5a_2} - \frac{A \sin(3t-3x)}{3a_1} + \sin(t-x) \right)^2 \\ &+ \sin(t+x) + \frac{A \sin(3t+3x)}{3a_1} + \frac{A \sin(5t+5x+\phi)}{5a_2} \end{aligned} \right)^2. \quad (7)$$

Figure 8 shows two examples of the energy functions computed analytically following Equation 7. In this particular example, the amplitude of the higher harmonic is maximum and the phase ϕ is 0 deg for the left image and 150 deg for the right image (note only half of a period is represented). When ϕ is equal to 0 deg, energy peaks are more uniformly distributed along the negative diagonal (direction from back to front corner in the image) than

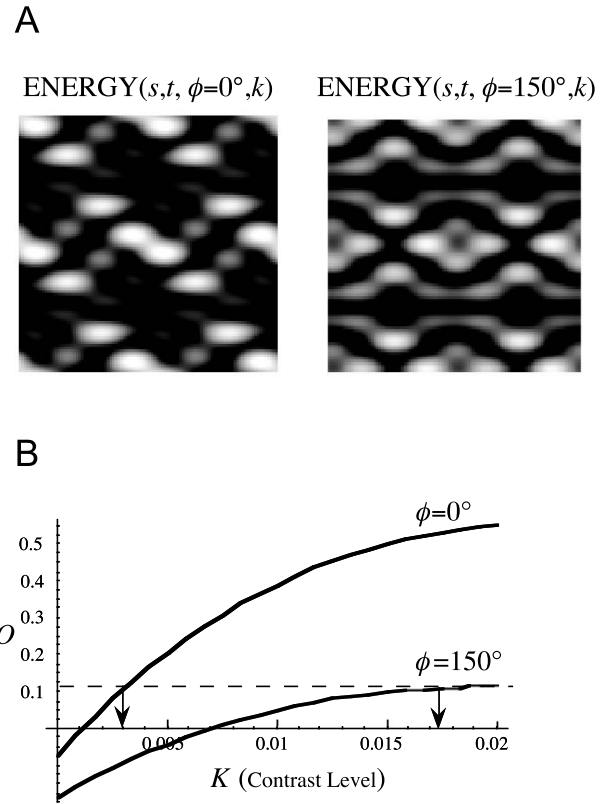


Figure 8. Procedure adopted to determine spatiotemporal orientation of maxima of local energy and to simulate quantitatively the contrast sensitivity of high harmonics to balance the motion direction bias. (A) Two examples of the energy functions for two values of the phase ϕ (0 and 150 deg), obtained by applying a broad quadrature phase filter ($a_1 = 0.36$ and $a_2 = 0.25$ in Equation 7). (B) Dependency of the motion contrast O on the contrast of the high harmonic plaid pattern $K = CA$. O is the output of the operator of Equation 4, obtained applying the convolution of the energy in A with the operators H_{left} and H_{right} , given by Equations 8 and 9, with σ_l and σ_e equal to 0.17 and 0.13 deg, respectively.

when ϕ is equal to 150 deg. If the symmetries in space-time of the energy ridges were to give the prevailing motion direction, we would predict a clear sensation of motion along this direction when ϕ is equal to 0 deg, but not when ϕ is equal to 150 deg. This qualitative prediction is confirmed by the experimental data of Figure 4. At the contrast of higher harmonic equal to 0.6, the stimulus at phase 0 is perceived to be drifting along the direction of the fifth harmonic, in phase with the fundamental (corresponding to -45 deg in the images of Figure 7), whereas no net motion was perceived for phase 150 deg.

To simulate quantitatively the data, we measured the motion energy contrast given by $O(\phi,K)$ as function of the contrast of the high harmonics. Figure 8B shows an example of a simulation for stimuli with phase equal to

0 and 150 deg and for a spatiotemporal profiles of second stage filters, given by difference of Gaussian distributions oriented along $x = \pm \frac{\kappa_0}{\omega_0} t = \pm 4t$ (see Figure 7C), that follow the equations:

$$H_{\text{left}}(x, t) = e^{-\frac{(x+4t)^2}{2\sigma_i^2}} \left(\frac{1}{\sigma_e} e^{-\frac{(x-4t)^2}{2\sigma_e^2}} - \frac{1}{\sigma_i} e^{-\frac{(x-4t)^2}{2\sigma_i^2}} \right) \text{ and} \quad (8)$$

$$H_{\text{right}}(x, t) = e^{-\frac{(x-4t)^2}{2\sigma_i^2}} \left(\frac{1}{\sigma_e} e^{-\frac{(x+4t)^2}{2\sigma_e^2}} - \frac{1}{\sigma_i} e^{-\frac{(x+4t)^2}{2\sigma_i^2}} \right), \quad (9)$$

where σ_i and σ_e , the space constants of the excitatory and inhibitory filters, are equal to 0.17 and 0.13 deg, respectively. The standard deviation along the orthogonal direction σ_l is always kept constant to 1 deg, which at the stimulus speed of 4 deg/sec corresponds to 0.25 s (the dependency of the model on the exact shape of the second stage filter is shown in Figure 9).

The dependency of O on the high frequencies contrast is not linear, given that an expansive nonlinearity is applied to the computation of the local energy function and a divisive normalizing term is used to estimate motion contrast (Equation 4). To simulate the psychophysical data, we imposed an arbitrary threshold on the motion energy contrast, chosen such to provide the best fit for the data corresponding to a phase shift of 90 deg (corresponding to a motion contrast of 0.1). For all the contrast curves at the various phase shifts, we evaluate the contrast of the higher harmonics corresponding to the same motion energy contrast threshold. Simulations are reported in Figure 9 for various choices of the free parameters. Figure 9A shows the simulation results calculated for each value of phase shift together with experimental data (black circles). In this case, the second stage filters given by Equations 8 and 9 were used. The different symbols refer to different shapes of front stage filters obtained by varying the relative gain of the various harmonics (parameters a_1 and a_2 in Equations 5 and 6). The factors correspond to filters with spatial bandwidth from 2.7 to 3.9 octaves and temporal bandwidth of 2.6–3.6 octaves, when the values at the three frequencies are fitted with a parabola on log units in spatial frequency (Morrone & Burr, 1988) and the temporal filter is considered very broad and flat in the temporal frequency domain. The model fits the data very well and it is robust to the shape of the front filters. It predicts also the fast decrease in sensitivity observed at $5\pi/6$, which is a peculiar characteristic of the perception of these stimuli.

We also simulated the data with different second stage filters, obtained with masks with different spatiotemporal distributions: with excitatory regions only (Figure 9B, diamond and circles); for two different sizes of the excitatory central Gaussian ($\sigma = 0.13$ and 0.07 deg); with excitatory and inhibitory regions (Figure 9B, square and triangles); for two different sizes of the excitatory central Gaussian ($\sigma = 0.13$ and 0.17 deg); and two different sizes of the inhibitory central Gaussian ($\sigma = 0.17$ and 0.3 deg).

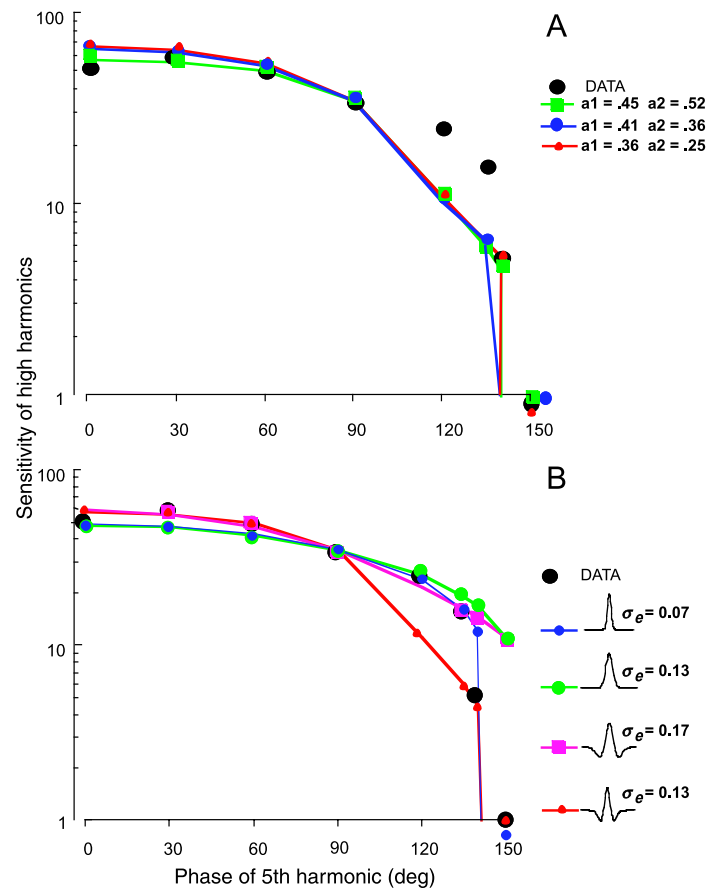


Figure 9. Dependency on simulation parameters. (A) Dependency on spatial and temporal bandwidth of front-end filters: blue ($a_1 = 0.41$, $a_2 = 0.36$); green ($a_1 = 0.45$, $a_2 = 0.52$); red ($a_1 = 0.36$, $a_2 = 0.25$). (B) Dependency on spatiotemporal profile and dimensions of space-time oriented filters at second stage: blue Gaussian mask ($\sigma = 0.07$ deg); green Gaussian mask ($\sigma = 0.13$ deg); magenta DoG mask ($\sigma_e = 0.17$ deg, $\sigma_i = 0.3$ deg); red DoG mask ($\sigma_e = 0.13$ deg, $\sigma_i = 0.17$ deg). Black circles represent experimental data for subject M.D.

The length of the mask (given by σ_l in Equations 8 and 9) was always kept constant. The results show that the second stage filters that best simulate the data are those with the smaller excitatory center. The presence of an inhibitory surround does not change substantially the pattern of results. The smaller filters are those that more closely match the size of the ridges of local energy.

Directionality of the front stage filter

The described algorithm used the same attenuation factors for each harmonic of the pairs with opposite motion directions. This corresponds to using a front stage filter that it is not directional but separable in space-time. This type of filter has been chosen because it optimizes the coding of phase between all components, without segregating

them for direction of motion. As evident from Equation 7, local energy from space-time separable filters contains terms at all possible beat frequencies between the various harmonics. A phase discrepancy between harmonics will generate beat frequencies with different phases, spreading the peaks of the energy function (for details and mathematical proof of the relation between local energy and phase congruency, see Morrone & Burr, 1988).

Here we modify the algorithm to study if the psychophysical data could still be simulated using two directional filters tuned to the two prevailing velocities, instead of a single nondirectional one.

Figure 10 shows the difference between the local energy calculated for the filter tuned to a velocity of 4 deg/s and that calculated for the filter tuned to the opposite velocity. Positive increasing numbers are represented with increasing value of red luminance, negative numbers with the green luminance. The Figure 10A shows the difference in the energy functions when the stimulus had a ϕ that was equal to 0 deg, Figure 10B when the stimulus had a ϕ that was equal to 150 deg. We measured the contrast of the harmonic required to reach threshold considering different manipulation of the two energy functions. The results are shown in Figure 10C. The scaling factors between the various curves are different and chosen arbitrarily to give the same value for the phase of 90 deg for an easy comparison between models. The contrast sensitivity did not vary with phase when the overall integral of the energy function for the filters tuned to positive and negative velocities are considered separately. This is a simple consequence of Parseval's theorem, given that the overall power of the output of the two filters is equal. However, the integral of the positive (orange symbols) and negative (green symbols) values of the difference between the outputs of the two directional filters also did not vary with phase. This indicates that an opponent stage mechanism is not sufficient per se to explain the dependency on phase. All of the above models would always predict perception of both direction of motion with no bias (remember that the scaling factor is arbitrary).

The introduction of a normalization stage (Georgeson & Scott-Samuel, 1999) after the motion opponent stage does not change significantly the predictions of the Adelson model because the normalization ($E_{\text{left}} + E_{\text{right}}$) is quite constant over space, and a model that considers or the maximum activity or the overall activity for the two directions also fails in simulating the phase dependency effect.

Introducing a second stage mechanism that measures elongation of the ridges of the output of the opponent stage did not improve the fitting of the data (purple curve). This model does show a dependence on phase, but the dependence is too modest to capture the pattern of the psychophysical data (black circles). For comparison, the results of the model of Figure 7 are also shown (red triangles).

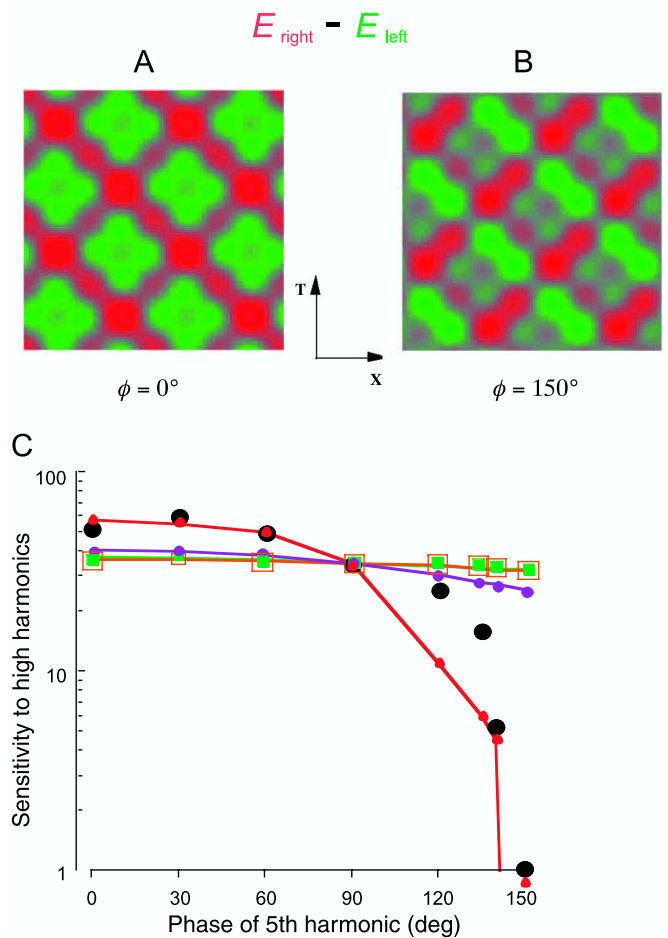


Figure 10. Comparison between models. (A and B) Energy obtained with a front-end rightward oriented filter minus energy obtained with a leftward oriented filter, respectively, when fifth harmonic is in phase with the others and when is phase shifted of 150 deg. (C) Experimental results (filled circles) and predictions of different models. Green curve: integral of positive part of $E_{\text{right}} - E_{\text{left}}$ (corresponding to rightward motion). Orange curve: integral of negative part of $E_{\text{right}} - E_{\text{left}}$ (corresponding to leftward motion). Purple curve: convolution between a DoG filter (oriented in the direction of motion Equations 8 and 9) and $E_{\text{right}} - E_{\text{left}}$. Red curve: output of the model proposed in this paper (see Figure 7).

Discussion

Here we report a new motion illusion, where the prevailing perceived direction of motion is strongly influenced by the relative phase of the harmonic components of the stimulus. The perception of two gratings that drift transparently in opposite directions can be biased towards a prevailing direction by changing the phase of one of the higher harmonics: the direction with greater phase congruency will prevail over the other direction. We also showed that the absolute phase of harmonic components is not an important parameter, excluding the possibility that

local luminance cues could be mediating the effect. It is not the phase shift of the single component that is important, but the relative relation between the phases of the various components (phase congruency).

Tolerance to phase discrepancy is quite large—up to about 90 deg—thereafter there is a rapid decrease in sensitivity. For phase differences between the two high harmonics of about 30 deg, the task becomes impossible, indicating that the neuronal underlying mechanisms are not able to distinguish this phase difference. Interestingly, spatial vision mechanisms seem to have a similar low sensitivity to phase both for tasks that require a simple form recognition (Bennett & Banks, 1987; Burr, 1980; Martini, Girard, Morrone, & Burr, 1996; Morrone, Burr, & Spinelli, 1989; Rentschler & Treutwein, 1985) and for tasks that require complex recognition of the general structure of transparent images (Morrone & Burr, 1997).

The analogy with spatial vision is quite interesting. The manipulation of the phase between harmonics of various orientations induces a perceptual bias in the prevailing orientation of the static scene (Morrone & Burr, 1997), even when the amplitude spectra are balanced between the various orientations. The similarity with the present findings indicates that the phenomenon is quite general. A similar mechanism probably mediates the sensitivity to phase congruency in space between different orientation bands and in space-time between different velocities and directions of motion. For static patterns, points of maximum phase congruency can be detected well by locating the local maxima in the local energy function (Morrone & Burr, 1988; Morrone & Owens, 1987), and these points usually correspond to salient features. The local energy maxima mark different types of visual features simultaneously, such as borders, specularities, shadows, bars, and combinations of them (Morrone & Burr, 1988, 1997). The organization of the feature map corresponds closely to the structure perceived by human observers and predicts many visual illusions (Morrone & Burr, 1997; Ross, Morrone, & Burr, 1989).

In the space-time domain, the velocity of features can be derived by evaluating the space-time orientation of the local energy ridges at each local maxima (Del Viva & Morrone, 1998; Zetzsche & Barth, 1991). These algorithms achieve both fine spatial localization and reliable estimation of velocity, fulfilling many of the demanding tasks imposed by our visual system (including the perception of non-Fourier stimuli). In these algorithms, the orientation of the ridge was evaluated by studying the local curvature, optimizing computational complexity. In other algorithms, phase congruency is explicitly evaluated and tracked over time (Fleet & Jepson, 1990). Here we evaluate the direction of the ridge by using second stage spatiotemporal filters of various orientations to simulate more closely the biological visual system. However, all these algorithms apply a second stage analysis after the computation of a local energy function.

The exact parameters of the front stage filters are not crucial, provided that the filter is sensitive to the frequency range between the first and the fifth harmonics. Filters that attenuate more the first harmonics than the fifth harmonics, over a range of 10 and 20 times, produce practically identical fitting results. However, to simulate adequately the data, all filters need to have separable spatiotemporal frequency tuning and hence be not directional. These kinds of filters are optimal to detect the phase congruency of the stimuli used here. Directional filters, that selectively sense only one of the two velocities will always perceive transparent balanced motion.

The selectivity of the second stage appears to be more important. Large filters, both with and without a center-surround inhibition, are less sensitive to the phase congruency and are not able to predict the rapid fall at about 120 deg phase shift. On the contrary, filters with a narrow spatiotemporal spread of the central region can fit the experimental data very well. The spread is about 5% of the periodicity of the input stimulus, indicating that the evaluation of the local curvature is very localized. We did not explicitly locate the peak and hence the features, but we simply measured the prevailing orientation of the output of the second stage filter. This strategy was sufficient to simulate the direction discrimination data. However, the simulation of more detailed motion perception may require that individual features are tracked and the orientation of the individual ridges need to be measured, as proposed by several other models (Del Viva & Morrone, 1998; Fleet & Jepson, 1990; Zetzsche & Barth, 1991).

There is ample evidence in the literature over the last two decades that the perception of motion is mediated by several different types of mechanisms that operate in parallel. Some aspect of motion requires only a first stage analysis (usually referred to as first order or Fourier motion) performed by directional filters (Adelson & Bergen, 1985; Burr, 1983; Burr & Ross, 1987; Watson & Ahumada, 1985). Others require a second order analysis (Badcock & Derrington, 1985; Cavanagh & Mather, 1989; Chubb & Sperling, 1988; Derrington & Badcock, 1985), usually implemented as an intrinsic nonlinearity applied at the input stage (Chubb & Sperling, 1988; Lu & Sperling, 1995, 2001). It is also widely accepted that if the power of the stimulus is not homogeneously distributed, the prevailing velocity will determine the saliency of perception (Chubb & Sperling, 1988; Georgeson & Scott-Samuel, 1999; Lu & Sperling, 1995; Zaidi & DeBonet, 2000). The stimuli used here are balanced in power and therefore the output of linear front stage mechanisms will be balanced. Even second order mechanisms based on motion energy could not explain the prevailing direction of motion and the dependence on phase (Adelson & Bergen, 1985; Adelson & Movshon, 1982; Movshon, Adelson, Gizzi, & Newsome, 1985). Second order mechanisms that compare locally outputs from different velocity would be insensitive to the phase parameter, although some of these algorithms

are very sophisticated and simulate many visual motion illusions successfully (Heeger, 1987; Weiss, Simoncelli, & Adelson, 2002; Yuille & Grzywacz, 1988) and some aspect of transparent motion perception (Qian et al., 1994a). In particular, an energy mechanism that locally measures the difference between opposite directions (Qian, Andersen, & Adelson, 1994b) can predict the transparency between simple random dot fields, but not the dependence of transparency on phase congruency, as shown by the red and green curves of Figure 10. To simulate the data, it is necessary that a second stage mechanism, oriented in space and time, follows the nonlinearity imposed by the front-end mechanism. This suggests that to simulate nonlinear perceptual phenomena like those illustrated here (and more generally the motion of contrast modulated stimuli), the spatial nonlinearity must precede the spatiotemporal correlation stage (in agreement with recent motion perception models; Benton, Johnston, McOwan, & Victor, 2001; Chubb & Sperling, 1988; Lu & Sperling, 2001; Solomon & Sperling, 1994; Turano & Pantle, 1989; Wilson et al., 1992). In this respect, the proposed model bears several similarities with the initial models proposed by Chubb and Sperling (1988) and Lu and Sperling (1995) and to the natural extension of this model where the motion is generated by features that belong to different domains like texture, color, and depth (Lu & Sperling, 2001, 2002). In all these models, it is the neuronal salience associated with the feature that is tracked over time. The standard model for luminance stimuli first performs a full-wave rectification after appropriate spatiotemporal separable filtering (called texture grabbing) and then applies standard Reichardt (1961) model to derive velocity (van Santen & Sperling, 1985). The computation of the local energy stage can perform the same function as feature grabbing. To distinguish which of the two models more closely simulates the neuronal mechanisms, specific tests need to be devised. However, the fact that motion did not vary with the global phase between harmonics favors the local energy alternative. The full-wave rectification of the texture grabbing modulus (Chubb & Sperling, 1988) would produce quite different outputs depending on the global phases of the stimuli. Global phase changes induce dramatic changes in luminance profiles and in the Michelson contrast of the present stimuli as much as a factor of 2. A full rectification would be highly sensitive to these variations.

The proposed model is similar to earlier edge detection models (Hildreth, 1984; Marr & Ullman, 1981) that compute feature velocity by tracking them over time.

The feature-tracking approach to the analysis of motion has been somewhat neglected over the last few decades, being considered biologically difficult to implement. However, some experimental evidence suggests that feature-tracking may be used by the human visual system (Cavanagh & Mather, 1989; Derrington & Ukkonen, 1999; Georgeson & Shackleton, 1989; Morgan, 1992; Morgan & Mather, 1994; Seiffert & Cavanagh, 1998; Turano &

Pantle, 1989) and that feature tracking may play an important role in solving the ambiguity of plaid stimuli (Alais, Wenderoth, & Burke, 1994, 1997; Bowns, 2002; Derrington & Ukkonen, 1999; Wilson et al., 1992). It is also important to note that several feature-tracking algorithms are able to mediate several types of first order motion. The present model would fail to detect the direction of motion of a simple sinusoidal grating given that it uses only the filter that is selective to zero speed. However, it has been developed to detect directional biases in energy-balanced stimuli and for these types of stimuli the stationary filter is the most selective. Analogously, for unbalanced motion energy stimuli, it is highly possible that the appropriate front stage energy filters would be the most active directional filters. In this case, the front-end nonlinear modulus should be substituted with a battery of motion energy filters tuned to different velocities. Current experiments on motion perception of compound gratings with different group velocities seem to support this hypothesis.

At present, it is difficult to assess whether a single motion detection mechanisms able to mediate all aspects of motion perception, or whether different mechanisms are necessary, and if these are totally independent (Clifford & Vaina, 1999; Smith & Ledgeway, 2001) or operate serially (Zanker, 1993). The present data indicate that one mechanism based on feature tracking applied after a nonlinear front stage could in principle handle and simulate several aspects of motion. The output soon after the computation of motion energy could be used to evaluate motion, being sensitive to Fourier motion and selective to high temporal frequencies. The feature tracking performed by the second stage directional mechanisms will provide a parallel evaluation of motion, selective to non-Fourier motion and less selective to high frequencies given the blur introduced by the second stage convolution. In this framework, the detection of Fourier and non-Fourier motion will have similar selectivity given the common front stage filtering. This is in agreement with several psychophysical results showing a common selectivity for first and second order motion (Benton et al., 2001; Smith & Ledgeway, 2001), but also with the known preference for lower temporal frequency for non-Fourier motion (Derrington, Badcock, & Henning, 1993; for review, see Lu & Sperling, 2001). It is also consistent with the finding of long-range motion mechanisms that operate at a high level (Braddick, 1980). The hypothesis is also consistent with several electrophysiological data that show directionally tuned mechanisms at the level of V1 and V2 that respond predominantly to the Fourier motion and with the properties of MT neurons showing a genuine selectivity for speed (Perrone & Thiele, 2001). Interestingly, the speed tuning of MT neurons is well simulated with the selectivity imposed by the second stage oriented mechanisms used to simulate the present data.

The predictive power of the model for the present data and the data about first and second order motion

characteristics from many laboratories indicate that the search for a mechanism that could handle many, if not all, aspects of motion perception should be pursued within the framework of feature-tracking models.

Acknowledgments

This work was supported by the PRIN Grant of the Italian Ministry of University and Research (MIUR).

Commercial relationships: none.

Corresponding author: M. Michela Del Viva.

Email: michela@in.cnr.it.

Address: Dipartimento di Psicologia, Università di Firenze, Firenze, Italy, & Istituto di Neuroscienze CNR, Pisa, Italy.

References

- Adelson, E. H., & Bergen, J. R. (1985). Spatiotemporal energy models for the perception of motion. *Journal of the Optical Society of America A*, 2(2), 284–299. [[PubMed](#)]
- Adelson, E. H., & Movshon, J. A. (1982). Phenomenal coherence of moving visual patterns. *Nature*, 300(5892), 523–525. [[PubMed](#)]
- Alais, D., Wenderoth, P., & Burke, D. (1994). The contribution of one-dimensional motion mechanisms to the perceived direction of drifting plaids and their after effects. *Vision Research*, 34(14), 1823–1834. [[PubMed](#)]
- Alais, D., Wenderoth, P., & Burke, D. (1997). The size and number of plaid blobs mediate the misperception of type-II plaid direction. *Vision Research*, 37(1), 143–150. [[PubMed](#)]
- Badcock, D. R., & Derrington, A. M. (1985). Detecting the displacement of periodic patterns. *Vision Research*, 25(9), 1253–1258. [[PubMed](#)]
- Bennett, P. J., & Banks, M. S. (1987). Sensitivity loss in odd-symmetric mechanisms and phase anomalies in peripheral vision. *Nature*, 326(6116), 873–876. [[PubMed](#)]
- Benton, C. P., Johnston, A., McOwan, P. W., & Victor, J. D. (2001). Computational modeling of non-Fourier motion: Further evidence for a single luminance-based mechanism. *Journal of the Optical Society of America A, Optics, Image Sciences, and Vision*, 18(9), 2204–2208. [[PubMed](#)]
- Bowns, L. (2002). Can spatio-temporal energy models of motion predict feature motion? *Vision Research*, 42(13), 1671–1681. [[PubMed](#)]
- Braddick, O. J. (1980). Low-level and high-level processes in apparent motion. *Philosophical Transactions of the Royal Society of London B, Biological Sciences*, 290(1038), 137–151. [[PubMed](#)]
- Burr, D. C. (1980). Sensitivity to spatial phase. *Vision Research*, 20(5), 391–396. [[PubMed](#)]
- Burr, D. C. (1983). Human vision in space and time. *Proceedings of the International Union of Physiological Sciences XV*, 510.504.
- Burr, D. C., & Ross, J. (1987). Visual analysis during motion. In M. A. Arbib & A. R. Hanson (Eds.), *Vision, Brain and Co-operative processes*. Boston: MIT Press.
- Cavanagh, P., & Mather, G. (1989). Motion: The long and short of it. *Spatial Vision*, 4(2–3), 103–129. [[PubMed](#)]
- Chubb, C., & Sperling, G. (1988). Drift-balanced random stimuli: A general basis for studying non-Fourier motion perception. *Journal of the Optical Society of America A*, 5(11), 1986–2007. [[PubMed](#)]
- Clifford, C. W., & Vaina, L. M. (1999). A computational model of selective deficits in first and second-order motion processing. *Vision Research*, 39(1), 113–130. [[PubMed](#)]
- Del Viva, M. M., & Morrone, M. C. (1998). Motion analysis by feature tracking. *Vision Research*, 38(22), 3633–3653. [[PubMed](#)]
- Derrington, A. M., & Badcock, D. R. (1985). Separate detectors for simple and complex grating patterns? *Vision Research*, 25(12), 1869–1878. [[PubMed](#)]
- Derrington, A. M., Badcock, D. R., & Henning, G. B. (1993). Discriminating the direction of second-order motion at short stimulus durations. *Vision Research*, 33(13), 1785–1794. [[PubMed](#)]
- Derrington, A. M., & Ukkonen, O. I. (1999). Second-order motion discrimination by feature-tracking. *Vision Research*, 39(8), 1465–1475. [[PubMed](#)]
- Fleet, D. J., & Jepsen, A. D. (1990). Computation of component image velocity from local phase information. *International Journal of Computerized Vision*, 5, 77–104.
- Georgeson, M. A., & Scott-Samuel, N. E. (1999). Motion contrast: A new metric for direction discrimination. *Vision Research*, 39(26), 4393–4402. [[PubMed](#)]
- Georgeson, M. A., & Shackleton, T. M. (1989). Monocular motion sensing, binocular motion perception. *Vision Research*, 29(11), 1511–1523. [[PubMed](#)]
- Grzywacz, N. M., & Yuille, A. L. (1990). A model for the estimate of local image velocity by cells in the visual cortex. *Proceedings of the Royal Society of London B, Biological Sciences*, 239(1295), 129–161. [[PubMed](#)]

- Heeger, D. J. (1987). Model for the extraction of image flow. *Journal of the Optical Society of America A*, 4(8), 1455–1471. [PubMed]
- Hildreth, E. C. (1984). The computation of the velocity field. *Proceedings of the Royal Society of London B*, 221, 189–220. [PubMed]
- Landy, M. S., Cohen, Y., & Sperling, G. (1984). HIPS: A Unix-based image processing system. *Computer Vision, Graphics and Image Processing*, 25, 331–347.
- Lu, Z. L., & Sperling, G. (1995). The functional architecture of human visual motion perception. *Vision Research*, 35(19), 2697–2722. [PubMed]
- Lu, Z. L., & Sperling, G. (2001). Three-systems theory of human visual motion perception: Review and update. *Journal of the Optical Society of America A, Optics, Image Science, and Vision*, 18(9), 2331–2370. [PubMed]
- Lu, Z. L., & Sperling, G. (2002). Stereomotion is processed by the third-order motion system: Reply to comment on “Three-systems theory of human visual motion perception: Review and update.” *Journal of the Optical Society of America A*, 19, 2144–2153.
- Marr, D., & Ullman, S. (1981). Directional selectivity and its use in early visual processing. *Proceedings of the Royal Society of London B, Biological Sciences*, 211(1183), 151–180. [PubMed]
- Martini, P., Girard, P., Morrone, M. C., & Burr, D. C. (1996). Sensitivity to spatial phase at equiluminance. *Vision Research*, 36(8), 1153–1162. [PubMed]
- Morgan, M. J. (1992). Spatial filtering precedes motion detection. *Nature*, 355(6358), 344–346. [PubMed]
- Morgan, M. J., & Mather, G. (1994). Motion discrimination in two-frame sequences with differing spatial frequency content. *Vision Research*, 34(2), 197–208. [PubMed]
- Morrone, M. C., & Burr, D. C. (1988). Feature detection in human vision: A phase-dependent energy model. *Proceedings of the Royal Society of London B, Biological Sciences*, 235(1280), 221–245. [PubMed]
- Morrone, M. C., & Burr, D. C. (1997). Capture and transparency in coarse quantized images. *Vision Research*, 37(18), 2609–2629. [PubMed]
- Morrone, M. C., Burr, D. C., & Spinelli, D. (1989). Discrimination of spatial phase in central and peripheral vision. *Vision Research*, 29(4), 433–445. [PubMed]
- Morrone, M. C., & Owens, R. (1987). Feature detection from local energy. *Pattern Recognition Letters*, 1, 103–113.
- Movshon, J. A., Adelson, E. H., Gizzi, M. S., & Newsome, W. T. (1985). *The analysis of moving visual patterns* (Vol. 54). Pontificiae Academiae Scientiarum Scripta Varia.
- Openheim, A. V., & Lim, J. S. (1981). The importance of phase in signals. *Proceedings of the IEEE*, 69, 529–541.
- Pantle, A., & Turano, K. (1992). Visual resolution of motion ambiguity with periodic luminance- and contrast-domain stimuli. *Vision Research*, 32(11), 2093–2106. [PubMed]
- Perrone, J. A., & Thiele, A. (2001). Speed skills: Measuring the visual speed analyzing properties of primate MT neurons. *Nature Neuroscience*, 4(5), 526–532. [PubMed]
- Piotrowski, L. N., & Campbell, F. W. (1982). A demonstration of the visual importance and flexibility of spatial-frequency amplitude and phase. *Perception*, 11(3), 337–346. [PubMed]
- Qian, N., Andersen, R. A., & Adelson, E. H. (1994a). Transparent motion perception as detection of unbalanced motion signals: I. Psychophysics. *Journal of Neuroscience*, 14(12), 7357–7366. [PubMed]
- Qian, N., Andersen, R. A., & Adelson, E. H. (1994b). Transparent motion perception as detection of unbalanced motion signals: III. Modeling. *Journal of Neuroscience*, 14(12), 7381–7392. [PubMed]
- Reichardt, W. (1961). Autocorrelation, a principle for evaluation of sensory information by the central nervous system. In W. Rosenblith (Ed.), *Sensory communications*. New York: Wiley.
- Rentschler, I., & Treutwein, B. (1985). Loss of spatial phase relationships in extrafoveal vision. *Nature*, 313(6000), 308–310. [PubMed]
- Ross, J., Morrone, M. C., & Burr, D. C. (1989). The conditions under which Mach bands are visible. *Vision Research*, 29(6), 699–715. [PubMed]
- Seiffert, A. E., & Cavanagh, P. (1998). Position displacement, not velocity, is the cue to motion detection of second-order stimuli. *Vision Research*, 38(22), 3569–3582. [PubMed]
- Smith, A. T., & Ledgeway, T. (2001). Motion detection in human vision: A unifying approach based on energy and features. *Proceedings of the Royal Society of London B, Biological Sciences*, 268(1479), 1889–1899. [PubMed]
- Solomon, J. A., & Sperling, G. (1994). Full-wave and half-wave rectification in second-order motion perception. *Vision Research*, 34(17), 2239–2257. [PubMed]
- Turano, K., & Pantle, A. (1989). On the mechanism that encodes the movement of contrast variations: Velocity discrimination. *Vision Research*, 29(2), 207–221. [PubMed]
- van Santen, J. P., & Sperling, G. (1985). Elaborated Reichardt detectors. *Journal of the Optical Society of America A*, 2(2), 300–321. [PubMed]

- Watson, A. B., & Ahumada, A. J., Jr. (1985). Model of human visual-motion sensing. *Journal of the Optical Society of America A*, 2(2), 322–341. [[PubMed](#)]
- Watson, A. B., & Pelli, D. G. (1983). QUEST: A Bayesian adaptive psychometric method. *Perception & Psychophysics*, 33(2), 113–120. [[PubMed](#)]
- Weiss, Y., Simoncelli, E. P., & Adelson, E. H. (2002). Motion illusions as optimal percepts. *Nature Neuroscience*, 5(6), 598–604. [[PubMed](#)]
- Wilson, H. R., Ferrera, V. P., & Yo, C. (1992). A psychophysically motivated model for two-dimensional motion perception. *Visual Neuroscience*, 9(1), 79–97. [[PubMed](#)]
- Yuille, A. L., & Grzywacz, N. M. (1988). A computational theory for the perception of coherent visual motion. *Nature*, 333(6168), 71–74. [[PubMed](#)]
- Zaidi, Q., & DeBonet, J. S. (2000). Motion energy versus position tracking: Spatial, temporal, and chromatic parameters. *Vision Research*, 40(26), 3613–3635. [[PubMed](#)]
- Zanker, J. M. (1993). Theta motion: A paradoxical stimulus to explore higher order motion extraction. *Vision Research*, 33(4), 553–569. [[PubMed](#)]
- Zetsche, C., & Barth, E. (1991). Direct detection of flow discontinuities by 3D-curvature operators. *Pattern Recognition Letters*, 12, 771–779.